

A Dynamic Theory of Optimal Capital Structure and Executive Compensation*

Andrew Atkeson

University of California, Los Angeles, Federal Reserve Bank of Minneapolis,
and NBER

Harold Cole

University of Pennsylvania and NBER

January 29, 2008

Abstract

We put forward a theory of the optimal capital structure of the firm based on Jensen's (1986) hypothesis that a firm's choice of capital structure is determined by a trade-off between agency costs and monitoring costs. We model this tradeoff dynamically. We assume that early on in the production process, outside investors face an information friction with respect to withdrawing funds from the firm that dissipates over time. We assume that they also face an agency friction that increases over time with respect to funds left inside the firm. The problem of determining the optimal capital structure of the firm as well as the optimal compensation of the manager is then a problem of choosing payments to outside investors and the manager at each stage of production to balance these two frictions. We show how this structure can generate a very rich theory of capital structure and compensation.

*We would like to thank Narayana Kocherlakota, Gian Luca Clementi, Willie Fuchs, Peter DeMarzo, Hugo Hopenhayn, Andy Skrzypacz, Steve Tadelis, Pierre-Olivier Weil. We thank Fatih Karahan for able research assistance. Atkeson gratefully acknowledge support from the National Science Foundation. Cole acknowledges the support of NSF SES 0137421.

1 Introduction

We put forward a theory of the optimal capital structure of the firm and the optimal compensation of the firm's managers based on Jensen's (1986) hypothesis that a firm's choice of capital structure is determined by a trade-off between agency costs and monitoring costs. We model this trade-off dynamically by assuming that outside investors in a firm face different obstacles to recouping their investment at different times. Early on in the production process, outside investors face an information friction — the output of the firm is private information to the manager of the firm unless the outside investors pay a fixed cost to monitor the firm. With time, the output of the firm is revealed to outside investors and, hence, the information friction disappears. At this later stage in the production process however, outside investors face an agency friction — the firm's manager can divert resources not paid out to investors in the early phases of production towards perquisites that provide him with private benefits. The stages correspond to subperiods within an information cycle that is repeated indefinitely. We associate this information cycle with an accounting or capital budgeting cycle at the firm. The problem of determining the optimal capital structure of the firm as well as the optimal compensation of the manager is then a problem of choosing payments to outside investors and the manager at each stage of production to balance these two frictions.¹

Our theory is developed in an dynamic optimal contracting framework, and, as a result, our model yields predictions about the joint dynamics of a firm's capital structure and its executive compensation. The choice of compensation for the manager is shaped by the assumption that the manager is risk averse while the outside investors are risk neutral. Our theory has the following implications regarding optimal capital structure and executive compensation. Each period, the payouts from the firm can be divided into payments to the manager that consist of a non-contingent base pay and a performance component of pay based on the realized output of the firm, as well as two distinct payments to the outside investors that resemble payments to debt and outside equity respectively. The debt-like payment to outside investors is made early in the period. It comes in the form of a fixed lump — the failure of which to pay leads to monitoring. The equity-like payment to outside investors comes in the form of a residual which depends upon the performance of the firm and is paid at the end of the period.

In our model, the fact that the manager receives some form of performance based pay

¹Our private information assumption is broadly consistent with Ravina and Sapienza (2006) that the trading behavior of executives and board members indicates the presence of substantial amounts of private information within firms.

is not motivated by the desire to induce the manager to exert greater effort or care in managing the firm. Instead, the performance based component of the manager’s pay simply serves to induce the manager to forsake expenditures on perquisites for his own enjoyment. Shocks that lead the agency friction to bind will lead to a performance bonus being paid, while negative shocks lead to the manager simply receiving his base pay. These results on compensation are consistent with the findings of the empirical literature, which shows that compensation is downwardly rigid, that good luck is rewarded, and that there is little empirical support for the relative compensation implication of the pure information based principal-agent model (e.g. Holmstrom 1982).

Since we also allow for productivity shocks which are publically observable at the beginning of the period, we can examine the impact of these observable shocks on compensation and capital structure as well. We find that base wages do not respond to observable productivity shocks in the optimal contract. We also find that the impact of positive observable productivity shocks on the agency friction and the performance bonus is dampened by increases in the extent of monitoring. When we alter the model to associate the observable shocks with managerial productivity, we show that the optimal retention strategy is to retain the manager if his productivity is above the threshold and fire him if he is below it. We also show that incumbent managers are protected against the risk that they become unproductive with a “golden parachute” as a direct consequence of optimal risk sharing.

We derive three dynamic results with respect to executive compensation. The first is that there is a simple monotonic relationship between the manager’s current total compensation and his future base wage. Moreover, when the manager and the outside investors share the same discount rate, the manager’s base wage tomorrow is equal to his total compensation today, and hence, his compensation is non-decreasing over time, regardless of the performance of the firm. This result is driven by the competing demands of consumption smoothing to minimize the costs of satisfying the promised utility constraint, and backloading of compensation to minimize the costs of satisfying the incentive constraint. We use this relationship to derive a simple recursive structure in terms of the base wage which allows to characterize the implications of our model.

The second dynamic result on compensation is that factors like the future growth prospects of the firm effect the agency friction with respect to the manager today through his conditional continuation utility. Fixing his base wage today, high future growth prospects increase his continuation utility conditional on his future base wage, and this relaxes the agency friction today. This relaxation then leads to a reduction in the likelihood of performance bonuses today. However, since their agency friction is more likely to bind in the future, it means a greater likelihood of performance bonuses and induced increases in base wages in the future.

The third dynamic result on compensation concerns the retention threshold when we associate the observable productivity shocks with managerial productivity. We show that the retention threshold is lower for managers who have been more productive in the past, and hence they are more likely to be retained. We interpret this result as a form of managerial entrenchment.

With respect to the link between executive compensation and capital structure, we find that the extent of the information and agency frictions that outside investors face depend crucially on the implicit utility or base wage level promised to the manager under the optimal contract. An increase in this promise relaxes the extent of the frictions, which leads to a reduction in monitoring and the current share of payments by the firm going to the debt holders. In this manner, the dynamics of executive compensation in our model drive the dynamics of the optimal capital structure. Positive productivity shocks lead to increase in firm profits. When these shocks cause the agency friction to bind today and hence lead to increases in future base wages, they thereby lead to reductions in the extent to which agency frictions bind in the future. This in turn leads to a reduction in the future level of monitoring and the share of payments going to debt. The downward rigidity of managerial compensation means that there is an important asymmetry in terms of the impact of positive vs. negative profitability shocks. Negative shocks (which do not lead to managerial turnover) do not effect future base wages and hence the only impact on capital structure is coming directly through the persistence of the profitability shock. Since anticipated current productivity shocks lead to an increase in likelihood of the agency friction binding, they result in an increase in current monitoring and the share of output going to pay debt holders. While future growth prospects, which come from high levels of future productivity, increase the conditional continuation payoff of the manager and hence reduce the likelihood that the current agency friction binds. This in turn reduces the extent of monitoring today and the share of payments going to debt.

Our theory also has implications for the relationship between the optimal financial structure of the firm and its optimal production plan. It predicts that there is a wedge between the marginal product of capital in the firm and rental rate of capital that depends upon the expected monitoring costs associated with bankruptcy and the inefficient risk-sharing between outside investors and the manager induced by the agency friction. The extent to which the agency frictions binds is also governed by the magnitude of the manager's base wage promise. Increases in base wages reduce the extent to which the friction binds, increase in the capital stock of the firm, and reduce the wedge between the internal and external return on capital. Under certain parametric assumptions, we are able to compute the magnitude of the wedge between the marginal product of capital and its rental rate in terms of readily

observed features of the firm’s financial structure and its executive compensation.²

Our dynamic model delivers predictions for the division of payments from the firm between the manager, the owners of outside equity, and the owners of the firm’s debt based on the trade-off of information and agency frictions. It is important to note that our dynamic model does not pin down the debt-equity ratio of the firm. This is because our model does not pin down the source of financing for ongoing investment in the firm. We conjecture that this failure of our model to pin down the debt-equity ratio of the firm in a dynamic setting may be a general feature of completely specified “trade-off” models of corporate finance.

We also use our model to examine the role of financial hedging in the firm’s optimal capital structure. In the data, firms are frequently seen to use financial instruments to hedge against both idiosyncratic and aggregate risks. According to standard theory, these financial hedges add no value. In our baseline model, financial hedging by the firm would actually be counter-productive. However, when we restrict ourselves to nonstate-contingent debt contracts, we show that hedges can play a role in fine-tuning the efficient contract in terms of achieving the optimal trade-off between bankruptcy risk and the agency friction.

This paper considers the optimal financial contract between outside investors and a manager in the presence of both information and agency frictions when there is the possibility of monitoring. It is therefore related to a wide range of prior research on each of these topics. The within period, or static, aspects of the information and monitoring aspect is similar to Townsend (1979), while the static aspects of the agency friction and the information friction are similar to Hart and Moore (1995), in that these frictions can rationalize a division of the firms output into debt, and other payments.³ However, unlike these prior papers, the inclusion of both frictions and monitoring, and the specific form of these friction leads to three different payment streams coming out of the firm, outside debt, outside equity, and managerial compensation.

Since we consider these frictions within a recursive environment, our paper is related to prior work on dynamic efficient contracting. However, unlike the literature on dynamic models of efficient financial contracting with information frictions, such as Atkeson (1991), Hopenhayn and Clementi (2002), Demarzo and Fishman (2004) or Wang (2004), our information friction is temporary since there is complete information revelation by the end of the period. As a result, while the costly state verification aspect of our model rationalizes outside

²While all of the models that generate debt constraints as part of the optimal contract generate a wedge between the inside return to capital and the outside cost of capital (e.g. Atkeson 1991, Hopenhayn and Clementi 2002, Albuquerque and Hopenhayn 2004, Bernake and Gertler 1989, and Charlstrom and Furest 1997), the advantage of our set-up is that it tightly ties this wedge to observable aspects of the contract.

³As in Jensen (1986) debt acts as a means of avoiding the agency friction associated with leaving funds in the firm and awaiting their payout as dividends.

debt, the dynamic aspects and the overall tractability of our model are similar to those of the dynamic enforcement constraint literature, such as Albuquerque and Hopenhayn (2004), and Cooley, Marimon and Quadrini (2004). In our model contracting is complete, subject to explicit information and enforcement frictions. This is in contrast to a large literature that seeks to explain various aspects of the financial structure of firms as arising from incomplete contracting.⁴

2 Model

Risk neutral outside investors contract with a risk averse entrepreneur to run a production technology in an infinite number of periods. Each of these periods are divided up into three subperiods. At the beginning of the first subperiod the production shock η , which is public information, is realized and capital K is supplied to the project. In the second subperiod the output level $y = \theta\eta F(K)$ is realized, however both y and the production shock θ are known only to the manager. The outside investors can monitor the output of the firm at cost $\gamma F(K)$. The investors can also request a payment v , which can be contingent both on the monitoring choice and the monitoring outcome. At the end of the second sub-period, the manager has the option of investing up to the fraction τ of the remaining output of the firm into perks that he consumes and otherwise he reinvests the remaining output of the firm at gross rate of return one.⁵ In the third sub-period, the realized value of the shock θ becomes public information, as well as the manager's division of the firm's output between perks and productive reinvestment. The manager is paid x in this third sub-period.

This production process is then repeated in subsequent periods. We interpret this cycle of information about production as corresponding to an accounting cycle or a capital budgeting cycle within the firm. For simplicity we will assume that θ is i.i.d. with expectation equal to 1, but we will allow η to be Markov. Since the capital and monitoring decisions are made after η is known, they can depend upon its realization. We assume that the rental rate on capital is r and that the outside investors discount the future at rate $1/R$.

⁴Examples include Hart and Moore (1995, 1997), as well as Aghion and Bolton (1992), which examines the efficient allocation of control rights, Dewatripont and Tirole (1994), in which outside investors choose their holdings of a debt as opposed to equity claims to generate the efficient decision with respect to the interference or not in the continuing operation of the firm, and Zweibel (1996), in which manager uses debt as a means of constraint their future investment choices to be more efficient.

⁵An alternative interpretation is that the manager has become essential to maintaining the value of the residual output in the third sub-period. Without his cooperation the value of this output is reduced by the factor $(1 - \tau)$ and that based upon this, the manager can, in the third subperiod, renegotiate his contract. For simplicity, we assume that the manager has all the bargaining power in this renegotiation, and hence he is able to demand that the fraction τ of the residual output be given to him.

We assume that all managers not running a project have an outside opportunity to enjoy consumption c_0 each period. Corresponding to this constant consumption flow is a reservation expected discounted utility level U_0 . Individual rationality requires that new managers can expect utility of at least U_0 under a contract and that incumbent managers can expect a utility of at least U_0 in the continuation of any contract.

We present a recursive characterization of the optimal dynamic contract. Because there is complete resolution of uncertainty at the end of each period, the persistence of the shocks does not generate any dynamic informational incompleteness in the model. Hence, a version of the revelation principal will apply here.⁶ Accordingly, we assume that the outside investor's contract with the incumbent manager is indexed by a utility level U promised him from this period forward and the prior realization of the public productivity shock η_{-1} because it is persistent. This utility level is a contractual state variable carried over from the previous period and hence is determined before the realization of the productivity shocks. We let $V(U, \eta_{-1})$ denote the expected discounted value of payments to outside investors given utility promise of U to the incumbent manager and the prior shock η_{-1} . We assume that the p.d.f. for η is given by $h(\eta|\eta_{-1})$, and the p.d.f. and c.d.f. for θ are given by $p(\theta)$ and $P(\theta)$ respectively.

A dynamic contract has the following elements. Given the utility U promised to the incumbent manager as a state variable, the contract specifies an amount of capital to be supplied to the project given the realized value of η , $K(\eta; U, \eta_{-1})$, a monitoring indicator function given η and the announcement $\hat{\theta}$, $m(\eta, \hat{\theta}; U, \eta_{-1})$, payments from the manager to the outside investors in the second subperiod $v_0(\eta, \hat{\theta}; U, \eta_{-1})$ if there is no monitoring and $v_1(\eta, \hat{\theta}, \theta; U, \eta_{-1})$ if there is monitoring, and payments from the outside investors to the manager in the third subperiod $x(\eta, \hat{\theta}, \theta; U, \eta_{-1})$. The recursive representation of the contract also specifies continuation utilities $Z(\eta, \hat{\theta}, \theta; U, \eta_{-1})$ for the incumbent manager. In what follows, we suppress reference to U and η_{-1} where there is no risk of confusion. Finally, we will find it useful to define M as the set of reports such that monitoring occurs; i.e. $M(\eta; U, \eta_{-1}) = \{\hat{\theta} : m(\eta, \hat{\theta}; U, \eta_{-1}) = 1\}$.

These terms of the contract are chosen subject to the limited liability constraints

$$v_0(\eta, \hat{\theta}) \leq \hat{\theta}\eta F(K(\eta)), \quad v_1(\eta, \hat{\theta}, \theta) \leq \theta\eta F(K(\eta)), \quad \text{and} \quad x(\eta, \hat{\theta}, \theta) \geq 0. \quad (1)$$

Since the incumbent manager can always quit and take his outside opportunity in the next

⁶This would be true even if the temporarily private shock θ was persistent since it is revealed in the third subperiod.

period, we have an individual rationality constraint

$$Z(\eta, \hat{\theta}, \theta) \geq U_0 \text{ for all } \eta, \hat{\theta}, \theta \quad (2)$$

We require that the contract deliver the promised utility U to the incumbent manager

$$\int_{\eta} \int_{\theta} [u(x(\eta, \theta, \theta)) + \beta Z(\eta, \theta, \theta)] p(\theta) h(\eta, \eta_{-1}) d\theta d\eta = U. \quad (3)$$

The incumbent manager must be induced to truthfully report θ in the second sub-period. For reports that don't lead to monitoring there is a required payment of $v_0(\eta, \hat{\theta})$, and the report $\hat{\theta} \notin M$ is feasible if $\theta \geq v_0(\eta, \hat{\theta})/(\eta F(K(\eta)))$.⁷ Hence we have incentive constraints, for all θ and $\hat{\theta} \notin M$ such that $\theta \geq v_0(\eta, \hat{\theta})/(\eta F(K(\eta)))$

$$u(x(\eta, \theta, \theta)) + \beta Z(\eta, \theta, \theta) \geq u(x(\eta, \hat{\theta}, \theta)) + \beta Z(\eta, \hat{\theta}, \theta). \quad (4)$$

Finally, there is a dynamic no-perks constraint arising from the assumption that the manager can spend fraction τ of whatever resources are left in the project at the end of the second sub-period on perks that he enjoys. This constraint is given by

$$\begin{aligned} u(x(\eta, \hat{\theta}, \theta)) + \beta Z(\eta, \hat{\theta}, \theta) &\geq u(\tau(\theta \eta F(K(\eta)) - v_1(\eta, \hat{\theta}, \theta))) + \beta U_0 \\ \text{if } m(\eta, \hat{\theta}) &= 1, \text{ and} \\ u(x(\eta, \hat{\theta}, \theta)) + \beta Z(\eta, \hat{\theta}, \theta) &\geq u(\tau(\theta F(K(\eta)) - v_0(\eta, \hat{\theta}))) + \beta U_0 \\ \text{o.w.} & \end{aligned} \quad (5)$$

for all η , θ and for all $\hat{\theta} \notin M$ such that $v_0(\eta, \hat{\theta}) \leq \theta \eta F(K(\eta))$. Here, in the left-hand side of (5), we have used the requirement that the manager's continuation utility $Z(\hat{\theta}, \theta)$ cannot be driven down below U_0 to compute the manager's utility in the event that he invests in perks and then is fired as a consequence.

The terms of the dynamic contract are chosen to maximize the expected discounted value of payments to the outside investors. This problem is to choose $K(\eta)$, $m(\hat{\theta})$, $v_0(\hat{\theta})$,

⁷Implicitly we're assuming that if the manager makes a report that doesn't lead to monitoring but doesn't pay $v_0(\eta, \hat{\theta})$, then he is monitored, his current consumption is set to 0 and his continuation to U_0 .

$v_1(\hat{\theta}, \theta)$, $x(\hat{\theta}, \theta)$, and $Z(\hat{\theta}, \theta)$ to maximize

$$V(U, \eta_{-1}) = \max_{\int_{\eta}} \left\{ \int_{\theta} \left\{ \begin{array}{l} (\theta\eta - \gamma m(\theta))F(K(\eta)) - x(\eta, \theta, \theta) \\ + \frac{1}{R}V(Z(\eta, \theta, \theta), \eta, \theta) \\ - rK(\eta) \end{array} \right\} p(\theta)d\theta \right\} h(\eta|\eta_{-1})d\eta \quad (6)$$

subject to the constraints (1), (2), (3), (4), and (5).

In the remainder of this section, we characterize elements of an efficient dynamic contract.

Proposition 2.1. *There is an efficient contract with the following properties: (i) $v_1(\eta, \hat{\theta}, \theta) = \theta\eta F(K(\eta))$ and $v_0(\eta, \hat{\theta}) = \theta^*(\eta)\eta F(K(\eta))$ where for each η , $\theta^*(\eta) = \inf \left\{ \hat{\theta} : m(\eta, \hat{\theta}) = 0 \right\}$ (ii) $x(\eta, \hat{\theta}, \theta) = 0$ and $Z(\eta, \hat{\theta}, \theta) = U_0$ for all η , and $\hat{\theta}$ such that $m(\eta, \hat{\theta}) = 1$ and $\hat{\theta} \neq \theta$ and $x(\eta, \hat{\theta}, \theta) = \tau(\theta - \theta^*(\eta))\eta F(K(\eta))$ and $Z(\eta, \hat{\theta}, \theta) = U_0$ for all η , and $\hat{\theta}$ such that $m(\eta, \hat{\theta}) = 0$ and $\hat{\theta} \neq \theta$, (iii) the set $\left\{ \hat{\theta} : m(\eta, \hat{\theta}) = 1 \right\}$ is an interval ranging from 0 to $\theta^*(\eta)$.*

Proof: Define $v_0^*(\eta) = \inf \left\{ v_0(\eta, \hat{\theta}) | m(\eta, \hat{\theta}) = 0 \right\}$ and $\theta^*(\eta) = \inf \left\{ \hat{\theta} | m(\eta, \hat{\theta}) = 0 \right\}$. Observe that to relax the constraint (4) as much as possible, the manager's utility following a misreporting of $\hat{\theta} \neq \theta$ should be set as low as possible. Given (1), (2), and (5), this gives $x(\eta, \hat{\theta}, \theta) = 0$, $Z(\eta, \hat{\theta}, \theta) = U_0$ for $\hat{\theta} \neq \theta$ when $m(\eta, \hat{\theta}) = 1$ (that is, when monitoring occurs), and

$$u(x(\eta, \hat{\theta}, \theta)) + \beta Z(\eta, \hat{\theta}, \theta) = u(\tau(\theta F(K(\eta)) - v_0(\eta, \hat{\theta}))) + \beta U_0$$

for $\hat{\theta} \neq \theta$ when $m(\eta, \hat{\theta}) = 0$. Note that given this, one never wants to misreport in a way that leads to monitoring. Also, note that the optimal misreport is $\hat{\theta}$ such that $v_0(\eta, \hat{\theta})$ is as small as possible or $v_0^*(\eta)$. Hence, we can combine the no-perks and the incentive constraint to get the fundamental constraint

$$u(x(\eta, \hat{\theta}, \theta)) + \beta Z(\eta, \hat{\theta}, \theta) = u(\tau(\theta F(K(\eta)) - v_0^*(\eta))) + \beta U_0 \quad (7)$$

Thus, this best possible report, $v_0^*(\eta)$, determines the extent to which the incentive and no-perks constraints bind. Holding fixed the monitoring set, setting $v_0^*(\eta)$ as high as is feasible relaxes these constraints as much as possible. Since feasibility requires that $\theta^*(\eta)\eta F(K(\eta)) \geq v_0^*(\eta)$, this gives us that under an optimal contract, $v_0^*(\eta) = \theta^*(\eta)\eta F(K(\eta))$.

That $\left\{ \hat{\theta} : m(\eta, \hat{\theta}) = 1 \right\}$ is an interval follows from the argument that including some $\theta > \theta^*$ in the monitoring set does nothing to relax (7) and does require resources for monitoring. That $x(\eta, \hat{\theta}, \theta) = \tau(\theta - \theta^*(\eta))\eta F(K(\eta))$ for all η , and $\hat{\theta}$ such that $m(\eta, \hat{\theta}) = 0$ & $\hat{\theta} \neq \theta$ follows from the result that $v_0^*(\eta) = \theta^*(\eta)\eta F(K(\eta))$. *Q.E.D.*

This proposition implies that our interim payment, v_i , shares the standard characteristics of a simple debt contract. If we interpret $\theta^*(\eta)\eta F(K(\eta))$ as the face value of the debt, then failure to pay this amount leads to monitoring, which we associate as bankruptcy, and the payment of everything to the creditors, while payment of $\theta^*(\eta)\eta F(K(\eta))$ means that no monitoring occurs.

This proposition implies that no one has an incentive to misreport since a report below $\theta^*(\eta)$ leads to monitoring and a report above $\theta^*(\eta)$ leads to no monitoring and an invariant intermediate payment. Hence, with this proposition, we can write our optimal contracting problem more simply as one of choosing capital $K(\eta)$, the upper support of the monitoring set $\theta^*(\eta)$, current managerial pay $w(\eta, \theta) = x(\eta, \theta, \theta)$, and continuation values $W(\eta, \theta) = Z(\eta, \theta, \theta)$ to maximize the payoff to the outside investors

$$V(U, \eta_{-1}) = \max_{\substack{w(\eta, \theta), W(\eta, \theta), \\ \theta^*(\eta), K(\eta)}} \int_{\eta} \left\{ \int_{\theta} \left\{ \begin{array}{l} \theta \eta F(K(\eta)) - w(\eta, \theta) \\ + \frac{1}{R} V(W(\eta, \theta), \eta) - rK(\eta) \\ - \gamma P(\theta^*(\eta) | \theta_{-1}) F(K(\eta)) \end{array} \right\} p(\theta) d\theta \right\} h(\eta | \eta_{-1}) d\eta \quad (8)$$

subject to the promise-keeping constraint

$$\int_{\eta} \int_{\theta} [u(w(\eta, \theta)) + \beta W(\eta, \theta)] p(\theta) h(\eta | \eta_{-1}) d\theta d\eta = U \quad (9)$$

and the no-perks constraint that for all η and all $\theta \geq \theta^*(\eta)$

$$u(w(\eta, \theta)) + \beta W(\eta, \theta) \geq u(\tau(\theta - \theta^*(\eta)) \eta F(K(\eta))) + \beta U_0. \quad (10)$$

If θ and η are bounded and $\beta R \leq 1$, then there will exist a utility level for the manager for which the no-perks constraint will not bind. Given this, we can bound the space of utility values for the manager and show that the recursive mapping defining V also satisfies the monotonicity and discounting conditions of Blackwell. Hence it is a contraction under the boundedness assumption.

The first-order conditions for this problem include

$$(\lambda + \delta(\eta, \theta)) u'(w(\eta, \theta)) = 1, \quad (11)$$

$$\frac{1}{R} V_1(W(\eta, \theta), \eta) + \beta(\lambda + \delta(\eta, \theta)) = 0, \quad (12)$$

$$\int_{\theta^*(\eta)}^{\infty} \delta(\eta, \theta) u'(\tau(\theta - \theta^*(\eta)) \eta F(K(\eta))) \tau \eta F(K(\eta)) d\theta = \gamma p(\theta^*) F(K(\eta)), \quad (13)$$

$$\left\{ \eta - \int_{\theta^*(\eta)}^{\infty} \delta(\eta, \theta) u'(\tau(\theta - \theta^*(\eta)) \eta F(K(\eta))) p(\theta) d\theta - \gamma P(\theta^*(\eta)) \right\} F'(K(\eta)) = r, \quad (14)$$

where λ is the multiplier on the promise-keeping constraint and $\delta(\eta, \theta)$ is the multiplier on the no-perks constraint.⁸ In addition, the envelope condition implies that

$$V_1(U, \eta_{-1}) = -\lambda.$$

Taken together, conditions (11) and (12) imply that when the no-perks constraint (10) doesn't bind, then $\delta(\eta, \theta) = 0$, and $w(\eta, \theta) = \bar{w}$, where

$$u'(\bar{w}) = 1/\lambda, \quad (15)$$

and that $w(\eta, \theta) \geq \bar{w}$, and strictly greater whenever the no-perks constraint binds. Since the rhs of (10) is increasing in θ , this implies that if there exists a $\bar{\theta}(\eta)$ such that it binds for all $\theta > \bar{\theta}(\eta)$, and the manager's payoff is strictly increasing in θ above $\bar{\theta}(\eta)$.

We will henceforth refer to \bar{w} as the *base wage*. We will refer to $w(\eta, \theta) - \bar{w}$ as the *performance bonus*. The key thing to note here is that the base wage is independent of η . The intuition for these results on compensation is that the ex ante marginal gain to increasing the manager's utility in state (η, θ) is $h(\eta|\eta_{-1})p(\theta)$, while the marginal cost of doing so is $\{1/u'(w(\eta, \theta))\} h(\eta|\eta_{-1})p(\theta)$. Efficiency therefore implies that $\min_{\theta} \{1/u'(w(\eta, \theta))\}$ is equalized for each η , and we get that the base wage is independent of η .

In addition, since (11) and (12) imply that

$$-\frac{1}{\beta R} V_1(W(\eta, \theta), \eta) u'(w(\eta, \theta)) = 1, \quad (16)$$

and hence, we get that

$$\frac{1}{\beta R} u'(w(\eta, \theta)) = u'(\bar{w}'),$$

where \bar{w}' denotes the base wage tomorrow when the no-perks constraint doesn't bind. This condition gives us a simple monotonic relationship $\omega(w)$ that characterizes the base wage tomorrow in terms of the wage rate today

$$\omega(w) \equiv u'^{-1}[u'(w)/\beta R], \quad (17)$$

⁸For simplicity of notation we have extended the definition of $\delta(\eta, \theta)$ to $\theta < \theta^*(\eta)$ and are simply taking it to be 0 for these values.

where $\omega' > 0$ and, when $\beta R = 1$, $\omega(w) = w$. This result implies that a binding no-perks constraint today triggers both an increase in compensation today in the form of a performance bonus, and an increase in future compensation in the form of an increase in the base wage rate.

The intuition for this result is that the marginal rate of transformation between the utility of the manager today and utility tomorrow, conditional on η and θ , is

$$MRT = \frac{1/u'(w(\eta, \theta))}{[1/u'(\bar{w}')] / R},$$

and the marginal rate of substitution is $1/\beta$. Equalizing these two gives us our wage updating equation (17).

Overall, efficient compensation is trading off, the desire to smooth compensation in order to minimize the total costs of satisfying the utility condition (9) against the desire to back load compensation in order to satisfy the enforcement constraint implied by (10) as costlessly as possible. This last effect arises because 1 unit of consumption today costs the investors as much as R units tomorrow, but the R units tomorrow help with the enforcement constraint both today and tomorrow.

We summarize our results on compensation in the following proposition.

Proposition 2.2. *Compensation comes in the form of a base wage \bar{w} , which is independent of η and θ , and a performance bonus $w(\eta, \theta) - \bar{w} \geq 0$ which is generated by the no-perks constraint (10) binding; triggered by a sufficiently high surprise profit shock ($\theta > \bar{\theta}(\eta)$). Increases in the current wage via a performance bonus lead to increases in future base wages according to (17). There will be an upward (downward) drift in base wages even without the no-perks constraint binding if βR is greater (less) than 1.*

The downward rigidity of compensation follows from our assumption of an enforcement friction in which the manager cannot be prevented from extracting a fraction of the residual output of the firm even if that extraction can subsequently be detected. This downward rigidity has been documented in the empirical literature on executive compensation. Tirole (2006) notes that managers tend to receive stable compensation despite poor performance. In addition, the implication that the manager's performance bonus is induced by sufficient positive shocks affecting firm profitability is also consistent with the empirical literature. Bertrand and Mullainathan (2001) find that managers are rewarded for luck, but not punished on the downside. A pure information friction would not have implied this sort of downward rigidity, and would also have implied that relative performance (the performance of the manager's firm relative to other firms which are likely to have been hit with correlated

shocks) would be an important factor in compensation. Tirole (2006) notes that relative performance is not used in executive incentive schemes (see also Jensen and Murphy 1990 or Barro and Barro 1990).

If η is i.i.d. then V depends solely on the promised utility of the manager, and when the constraint (10) binds, the optimal choices of $w(\theta)$ and $W(\theta)$ satisfy

$$1 = -\frac{1}{\beta R} V'(W(\theta)) u'(w(\theta))$$

and (10) as an equality. Hence, $w(\theta)$ and $W(\theta)$ are both increasing in θ when this constraint binds. Since we know from our wage updating condition (17) that $w(\theta)$ and next period's base wage, \bar{w}' , are monotonically related, this implies that \bar{w}' and $W(\theta)$ are also monotonically related. We also know from the envelope condition $V_1(W(\eta, \theta)) = -1/u'(\bar{w}')$. Thus, an increase in the future base wage implies an increase in the continuation utility of the manager, and decreases the continuation payoff of the investors. Because we have not been able to sign V_{12} in the non-i.i.d. case, we have not been able to prove this result more generally, though it is intuitive that it will hold. This intuition is consistent with the numerical examples we present below.

Proposition 2.3. *If η is i.i.d. an increase in tomorrow's base wage implies an increase in continuation utility of the manager and a decrease in the continuation payoff of the investors.*

The standard result due to Modigliani and Miller (1958) is that in a frictionless world, the capital structure of a firm has no impact on its efficient production plan. If the monitoring cost $\gamma = 0$, the optimal contract specifies that the outside investors monitor the output of the project in the second sub-period for all values of $\hat{\theta}$ and pay the manager constant compensation \bar{w} independent of the realized value of θ . In this environment the efficient capital stock satisfies $\eta F'(K) = r$ since the expected value of θ is one. Hence, we refer to an economy in which the monitoring cost $\gamma = 0$ as a *frictionless environment*. In contrast, with financial frictions, there is a wedge between the marginal product of capital and its rental rate. From the first-order condition for capital (14), one can directly deduce the following proposition.

Proposition 2.4. *If either (i) $\gamma > 0$ and monitoring occurs with positive probability or (ii) the no-perks constraint binds with positive probability, then*

$$\eta F'(K(\eta)) < r.$$

To gain greater insight into this wedge, we use conditions (11) and (14), to get that

$$\left\{ - \int_{\theta^*(\eta)}^{\infty} \left\{ \left[\frac{u'[\tau(\theta - \theta^*(\eta))\eta F(K(\eta))]}{u'(w(\eta, \theta))} - \frac{u'[\tau(\theta - \theta^*(\eta))\eta F(K(\eta))]}{u'(\bar{w})} \right] \tau(\theta - \theta^*(\eta))\eta p(\theta) \right\} d\theta \right\} F'(K(\eta)) = r. \quad (18)$$

From this expression, one can see that there are two parts to this wedge between $\eta F'(K)$ and r . The first part, $\gamma P(\theta^*(\eta))$, is the expected loss due to monitoring. This loss is a cost of debt since the monitoring that debt requires in the event of monitoring results in a loss of output. The second part of the wedge is the loss due to inefficient risk-sharing between the outside investors and the manager that arises as a result of the performance based component of the manager's compensation (i.e. to the extent that $u'(w(\eta, \theta)) < u'(\bar{w})$). Specifically, this is the loss due to the fact that the risk averse manager places a lower valuation on the state contingent component of his compensation than the outside investors do. If these costs are positive, then the level of investment is low relative to the frictionless environment.

Condition (13) implies that under the efficient contract, $\theta^*(\eta)$ is determined by a trade-off between the marginal cost of monitoring as captured by the right hand side of this expression, and the marginal impact of monitoring on the cost of distorting the manager's consumption, as captured by the left hand side of this expression. Besides determining the face value of the debt, the choice of $\theta^*(\eta)$ determines the share of gross output going to debt, which is the inverse of the interest coverage. This share is given by

$$\begin{aligned} & \frac{\int_0^{\theta^*(\eta)} \theta \eta F(K(\eta)) p(\theta) d\theta + (1 - P(\theta^*(\eta))) \theta^* \eta F(K(\eta))}{\eta F(K(\eta))} \\ &= \int_0^{\theta^*(\eta)} \theta p(\theta) d\theta + (1 - P(\theta^*(\eta))) \theta^*(\eta), \end{aligned} \quad (19)$$

and hence this share is monotonically increasing in $\theta^*(\eta)$. We focus on this measure of the magnitude of relative debt because, as we discuss later, the predictions for debt vs. equity are less precise, and this measure of leverage is more relevant for the issue of monitoring and transferring control of the firm from equity holders to debt holders (see Rajan and Zingales 1995 for a similar argument).

2.1 Characterization and Comparative Statics

Our results on compensation allow a simple recursive characterization of the optimal contract in terms of the base wage. Because the base wage is determined by the prior period's compensation level, and since the optimal conditions for monitoring and capital are essen-

tially static, this characterization will enable us to derive comparative statics results with respect to the base wage with making assumptions about the stochastic process for η . Note that this is occurring despite the fact that we have not signed V_{12} .

Let $\Omega(\bar{w}, \eta_{-1})$ denote the manager's continuation utility in terms of the base wage \bar{w} and last period's public shock η_{-1} . Then the wage function today is simply the maximum of the base wage today and the wage that satisfies the no-perks constraint, or

$$w(\eta, \theta) = \max[\bar{w}, \tilde{w}(\eta, \theta)] \quad (20)$$

where $\tilde{w}(\eta, \theta)$ is such that

$$u(\tilde{w}(\eta, \theta)) + \beta\Omega(\omega(\tilde{w}(\eta, \theta)); \eta) = u(\tau(\theta - \theta^*(\eta))\eta F(K(\eta))) + \beta U_0. \quad (21)$$

Given this optimum wage function, we can use the first-order condition (11) to determine $\delta(\eta, \theta)$ and hence reduce the first-order conditions for the optimum level of monitoring $\theta^*(\eta)$ (13) and capital $K(\eta)$ (14) to a pair of simultaneous equations

$$\eta\tau \int_{\theta^*(\eta)}^{\infty} \left(\frac{u'[\tau(\theta - \theta^*(\eta))\eta F(K(\eta))]}{u'(w(\eta, \theta))} - \frac{u'[\tau(\theta - \theta^*(\eta))\eta F(K(\eta))]}{u'(\bar{w})} \right) p(\theta) d\theta = \gamma P(\theta^*), \quad (22)$$

and

$$\left\{ - \int_{\theta^*(\eta)}^{\infty} \left\{ \begin{array}{c} [\eta - \gamma P(\theta^*(\eta))] \\ \left[\frac{u'[\tau(\theta - \theta^*(\eta))\eta F(K(\eta))]}{u'(w(\eta, \theta))} - \frac{u'[\tau(\theta - \theta^*(\eta))\eta F(K(\eta))]}{u'(\bar{w})} \right] \times \\ \tau(\theta - \theta^*(\eta))\eta p(\theta) \end{array} \right\} d\theta \right\} F'(K(\eta)) = r, \quad (23)$$

which can be solved directly for the monitoring and capital choices. Finally, given the wage function, monitoring and capital choices, we can recursively define the manager's and the investors payoff as

$$\Omega(\bar{w}; \eta_{-1}) = E \left\{ u(w(\eta, \theta)) + \beta\Omega(\omega(w(\eta, \theta)); \eta) \mid \eta_{-1} \right\}, \quad (24)$$

and

$$\Phi(\bar{w}; \eta_{-1}) = \int_{\eta} \left\{ \begin{array}{c} [\eta - \gamma P(\theta^*(\eta))] F(K(\eta)) - rK(\eta) \\ + \int_{\theta} \left\{ \frac{1}{R} \Phi(\omega(w(\eta, \theta)); \eta) - w(\eta, \theta) \right\} p(\theta) d\theta \end{array} \right\} h(\eta \mid \eta_{-1}) d\eta, \quad (25)$$

where $\Phi(\bar{w}; \eta_{-1})$ denotes the investor's payoff.

With this recursive structure we can prove the following proposition about comparative statics results for $\theta^*(\eta)$ and $K(\eta)$.

Proposition 2.5. *If we are at an interior optimum, then fixing $K(\eta)$, $d\theta^*(\eta)/d\bar{w} < 0$ and $d\theta^*(\eta)/d\gamma < 0$, and fixing $\theta^*(\eta)$, $dK(\eta)/d\bar{w} > 0$.*

Proof: See the Appendix.⁹

Proposition 2.5 indicates that there is a natural sense in which \bar{w} is governing the extent of the overall agency friction within the model, and that the extent of monitoring and the level of capital are both increasing in \bar{w} , fixing the other, since increases in \bar{w} decrease the extent of this friction. It is trivial to show that in a larger sense monitoring becomes less frequent and capital rises towards its frictionless efficient level as the base wage becomes large. To pick an extreme example, if the base wage was so large that the no-perks constraint could never bind even when no monitoring is occurring, then it must be the case that monitoring is zero and the level of capital satisfies the frictionless efficiency condition $\eta F'(K(\eta)) = r$.

We have shown that monitoring decreases when the cost increases, fixing the level of capital. It is natural to suspect that it increases if we increase the agency friction by increasing τ . While comparative statics results with respect to the impact of τ are in general quite messy, the special case in which $\beta = 0$ delivers a very simple form for the optimal monitoring condition, and a straightforward comparative statics results with respect to the impact of changes in τ on θ^* . When $\beta = 0$, the first-optimality condition becomes

$$\eta\tau \int_{\theta^*(\eta)}^{\infty} \left(1 - \frac{u'[\tau(\theta - \theta^*)\eta F(K(\eta))]}{u'(\bar{w})} \right) p(\theta) d\theta = \gamma p(\theta^*),$$

Proposition 2.6. *If $\beta = 0$ and we are at an interior optimum, then $d\theta^*(\eta)/d\tau > 0$*

Proof: The derivative of the lhs of the above expression with respect to τ is positive. At an interior optimum, the second derivative with respect to $\theta^*(\eta)$ is negative and hence the results follows. *Q.E.D.*

It is interesting to note here that both increases in γ and increases in τ increase the extent to which the no-perks constraint binds. However, they move the capital structure of the firm in opposite directions. An increase in the friction coming from an increase in bankruptcy costs lowers the share of output going to debt, while an increase in the friction coming from an increase in the fraction that a manager can appropriate increases the share of output going to debt. The prediction that increases in γ lower our measure of leverage is consistent with Rajan and Zingales's (1995) finding that internationally leverage is negatively associated with stricter bankruptcy laws if we interpret strictness as implying higher costs.

⁹In the Appendix we discuss why a more general result is not possible, and hence we must fix K and θ^* in doing our comparative statics analysis of θ^* and K respectively.

2.2 Dynamics

To illustrate the dynamic implications of our model, we will examine several special cases using both analytic and numerical results. In our numerical examples we assume that the manager has log preferences, that $\beta = .75$, $\tau = .5$, and $\gamma = .25$. We assume that managers and investors have identical discount rates, which implies that the current total wage (base wage plus performance bonus) will equal tomorrow's base wage according to (17). The production function is given by $K^{\cdot 6}$ and the rental price is 1. For the shock θ we will assume that it is an independently distributed log-normal random variable with $\log(\theta) \sim N(-0.3/2, 0.3)$. For the public shock η , we will consider several cases. The numerical example is interesting both in terms of illustrating the workings of the model and because in this case we can establish that $\Omega(\bar{w}; \eta_{-1})$ is increasing in η_{-1} and this generates additional insights.

No Public Shocks: The first case we want to examine is where $\eta = 1$ forever. In this case, the dynamics of our model are coming solely through the effects of θ shocks on the no-perks constraint, and the resulting increase in continuation base wages. Since $\Omega(\bar{w}; \eta_{-1})$ is independent of η_{-1} , and $\Omega(\bar{w})$ is increasing in \bar{w} from proposition 2.3. Our wage equation (21) determining $\tilde{w}(\theta)$ becomes simply

$$u(\tilde{w}(\theta)) + \beta\Omega(\omega(\tilde{w}(\theta))) = u(\tau(\theta - \theta^*)F(K)) + \beta U_0.$$

We can define $\bar{\theta}$ by the requirement that

$$u(\bar{w}) + \beta\Omega(\bar{w}) = u(\tau(\bar{\theta} - \theta^*)F(K)) + \beta U_0.$$

Then, it follows that the no-perks constraint binds for values of θ for which $\theta > \bar{\theta}$. For realizations of $\theta > \bar{\theta}$, the need to deter rent grabbing by the manager leads to an increase in his compensation both today, and his future base wage, which in turn leads to an increase in his total payoff. The payoff to the investors is higher today, but lower tomorrow because of the increase in the promised level of future compensation.

Figure 1 illustrates these aspects of the optimal contract using several plots. The optimum level of monitoring in panel 1 is decreasing in the base wage, and goes to zero when the likelihood of the no-perks constraint binding goes to zero. The optimum level of the capital stock in panel 2 rises monotonically to its no-frictions efficient level as the base wage increases. The third panel plots the wage function $w(\theta)$ for a particular value of the base wage. The wage function is flat at the base wage up to a sufficiently high value of θ that the no-perks constraint binds, and then rises monotonically thereafter.

To understand the dynamic implications of this figure, assume that we're initially at the

base wage chosen in panel 3. Then, if θ is sufficiently low that the no-perks constraint doesn't bind, the base wage tomorrow will be the same as today. In which case, monitoring and the capital level will also be the same. At this level of the base wage, θ^* is about 0.45 and the capital stock is 89% of the frictionless efficient level. If θ is sufficiently high - above 0.93 - then the no-perks constraint will bind and both the current wage and the future base wage will need to be higher in order to satisfy the no-perks constraint. This rise in the current wage will mean that the manager earned a performance bonus this period. The resulting rise in the future base wage will imply a lower level of monitoring tomorrow and a higher level of capital.

Our model therefore implies that unanticipated profitability shocks, if they are sufficient large, will lead to a decrease in the share of output going to debt and hence an increase in the interest coverage. This prediction is consistent with the finding reported in Rajan and Zingales (1995) that profitability and leverage are negatively correlated. These shocks will also lead to an increase in the capital stock of the firm and a performance bonus being paid to the manager along with his base wage. On the other hand, negative unanticipated profitability shocks leave the share of output going to debt and the capital stock unchanged, and imply that the manager will receive only his base wage.

Future Growth Prospects: The second case that we want to consider is one in which we compare what happens under two different scenarios with respect to future growth prospects. In the first scenario η today is 1, but will be 3 from tomorrow onwards, while in the second scenario η is again 1 forever. We label the first scenario growth and the second no-growth, and Figure 2 present several key elements of the efficient contract. The first panel presents the efficient level of monitoring under the two scenarios. The growth case has much lower monitoring levels than the no-growth case for all base wages in which monitoring is positive. The efficient levels of the capital stock reverse this monitoring pattern, with the efficient level of the capital stock in the growth case lying above the no-growth level for all base wages that induce capital levels below the efficient level. The reason this is occurring can be gained from the second panel, which shows that the continuation payoff for the manager in the growth case lies weakly above the continuation payoff in the no-growth case. The fact that the continuation payoff lies above means that the no-perks constraint is less binding in the growth case than in the no-growth case, and this leads to monitoring being lower and the optimal capital stock being higher. This is also exhibited in the third panel where for a particular base wage we have plotted the two wage functions. The growth wage function is constant at the base wage for almost all of the θ values we consider, while the no-growth wage function rises monotonically starting at θ around 0.93. To understand why the continuation payoff for the manager is weakly higher in the growth case, note that this pattern on the

wage functions is reversed next period, where now in the growth case $\eta = 3$ and the wage function of the manager at this same base wage begins to rise at θ around 0.73 and lies strictly above the no-growth wage function for all θ 's above this level. Finally, note that the overall payoff to the manager in the growth case is higher conditional on his base wage, and hence, if we were equalizing initial payoffs to the manager, the first period base wage would be lower in the growth scenario than the no-growth scenario.

This example has illustrated how the impact of the future comes in through the continuation payoff of the manager, which is a fundamental feature of this model. It also illustrates how shifts up in the continuation function $\Omega(\bar{w})$ leads to a lower level of monitoring, a higher level of capital, and a reduced likelihood of the manager earning a performance bonus. This implies that firms that are anticipated to grow rapidly, such as many small firms, will have a low share of output going to debt and a high interest rate coverage. In the data, the ratio of the market-to-book value of assets is often taken to be a positive predictor of future growth prospects. This ratio is negatively correlated with leverage in the data, which is consistent with this prediction of the model (see Rajan and Zingales 1995). In the model, much of the manager's total compensation is backloaded and will come in the form of future performance bonuses and their subsequent impact on his future base wages.

I.I.D. Shocks: Here we want to examine the dynamic implications of our model for the case where the η shocks are i.i.d.. In this case, $\Omega(\bar{w}; \eta_{-1})$ is independent of η_{-1} . Our wage equation (21) determining $\tilde{w}(\eta, \theta)$ becomes

$$u(\tilde{w}(\eta, \theta)) + \beta\Omega(\omega(\tilde{w}(\eta, \theta))) = u(\tau(\theta - \theta^*(\eta))\eta F(K(\eta))) + \beta U_0,$$

and we can define $\bar{\theta}(\eta)$ by the requirement that

$$u(\bar{w}) + \beta\Omega(\bar{w}) = u(\tau(\bar{\theta}(\eta) - \theta^*(\eta))\eta F(K(\eta))) + \beta U_0.$$

Then, for each η , we have shown that the no-perks constraint binds for high values of θ for which $\theta > \bar{\theta}(\eta)$.

To further illustrate the working of the i.i.d. case we also computed a numerical example, where the magnitude of the shocks is chosen to capture something like normal cyclicalty rather than growth so we will assume that η takes on the values $\eta^h = 1.15$ and $\eta^l = 0.85$ with equal likelihood. Figure 3 plots several variables associated this case. In the first panel we display the optimal monitoring levels for both shocks. The optimal monitoring levels differ fairly sharply with respect to η , with higher levels of η being associated with weakly higher levels of monitoring. Similarly, the second panel also shows that the optimal capital

level differs sharply with respect to η , with higher η 's being associated with bigger levels of capital. The reason for this becomes clear in the third panel where the wage function has been plotted for each η . High η 's lead to a weakly higher wage level, and hence a tighter no-perks constraint, which in turn implies that a higher level of monitoring is optimal.

These results highlight the model's very different predictions for anticipated profitability shocks than for unanticipated profitability shocks. Anticipated shocks lead to an increase in size (here $F(K(\eta))$), the stock of fixed assets ($K(\eta)$), and the current share of output going to debt and hence a decrease in current interest coverage. While the current version of the model implies that in the long-run the share of output going to debt goes to zero, when we consider the optimal retention problem for the manager, this implication will no longer be true.

I.I.D. vs. Persistent Shocks: Here we wanted to do one final comparison in which we consider the implications of our model under two different scenarios. The first is simply the i.i.d. case that we just considered, while the second differs only by the assumption that the shocks are more persistent. We assume in the persistent shock case that the transition matrix is symmetric with probably 0.8 the value of tomorrow's η is unchanged from today. Figure 4 plots several variables from these two scenarios. Here again we see in panel 1 that the optimal monitoring levels are increasing in η , and that the persistent shock outcomes are not as extreme in their variation with η as with the i.i.d. shock, but this difference is very small. The second panel shows the wage functions for the two η cases and the two scenarios. The wage functions conditional on η are quite similar. The third panel shows the continuation payoffs for the manager, and just as in the growth cases, a persistent η shock implies a higher level of the continuation payoff in the high η case and a lower level in the low η case. The upward shift in the continuation payoff function $\Omega(\bar{w}, \eta_{-1})$ as a consequence of the shift in η_{-1} , raises the lhs of (21), but there is also a shift up in the current amount that can be taken in perks, fixing the monitoring threshold θ^* , which raises rhs of (21). These two effects are in an offsetting direction and quantitatively the shifts turn out to be small. As a result, the difference in the share of output going to debt is quite small between the two scenarios, conditional on η . However, there are large differences in the payoff to the investors across the two scenarios. For example at the same base wage as we used in the panel 2, the ratio of the conditional payoff to the investors given η^h relative to it given η^l is 1.24 in the i.i.d. case and 1.48 in the persistent case.

This finding that the interest share going to debt and interest coverage are very similar across scenarios which lead to substantial differences in the payoff to investors is interesting in light of the empirical results reported in Welch (2004). Welch reports that firms do little to offset changes in the impact of the market price of their equity on the debt-to-equity ratio,

and that as a result this ratio varies closely with stock prices. In our model, the fact that firms do not respond very differently to moderate profitability shocks depending on whether they are temporary or persistent will imply that the present value of the payoffs to equities will vary substantially without much change in our measure of the capital structure, the share of output going to debt.

2.3 Retention, Firing and Golden Parachutes

Thus far, in our dynamic model, in equilibrium the incumbent manager is never fired. We now extend our dynamic model to include a decision about whether to retain the incumbent manager. To do so, we consider an extension of the model in which we associate η with the current manager. The investors now have an incentive to retain incumbent managers with high productivity, or η , and replace those with low productivity. To keep things simple, we will assume that η draws are i.i.d. over time and that new managers start with $\eta_0 = 1$ and have reservation utility U_0 .

As in the basic model, the outside investors are deciding how to compensate the manager across realizations of his observable productivity η , but in addition, they are also deciding for which values of η they are going to retain the manager. We will assume that in the event the manager is not retained, his future continuation level is given by U_0 , but his current consumption is determined by the compensation offered him under his contract with the outside investors, w^F .

The outside investors problem of determining the optimal contract is separable into a two-stage contracting problem in which the outside investors first determine the allocation of utility across states and retention, and then determine the conditional optimal contract.

Stage 1: Decide whether to retain the manager and how to allocate utility conditional on η

$$V(U) = \max_{\substack{\delta(\eta) \in [0,1] \\ U^R(\eta), w^F(\eta)}} \int_{\eta} \left\{ \begin{array}{l} \delta(\eta) V^R(U^R(\eta), \eta) \\ + (1 - \delta(\eta)) [V^R(U_0, 1) - w^F(\eta)] \end{array} \right\} h(\eta) d\eta,$$

subject to

$$\int_{\eta} \{ \delta(\eta) U^R(\eta) + (1 - \delta(\eta)) [u(w^F(\eta)) + \beta U_0] \} h(\eta) d\eta.$$

For future reference, note that our f.o.c.'s include

$$-V_1^R(U^R(\eta), \eta) = \omega,$$

$$u'(w^F(\eta)) = 1/\omega.$$

Stage 2: Determine optimal compensation, capital and monitoring given η and the utility of the retained/new manager being U :

$$V^R(U, \eta) = \max \int \left\{ \theta \eta F(K(\eta)) - w(\eta, \theta) + \frac{1}{R} V(W(\eta, \theta)) \right\} p(\theta) d\theta \\ - \gamma P(\theta^*(\eta)) \eta F(K(\eta)) - rK(\eta)$$

subject to the promise-keeping constraint

$$\int [u(w(\eta, \theta)) + \beta W(\eta, \theta)] p(\theta) d\theta = U$$

and the dynamic no-perks constraint

$$u(w(\eta, \theta)) + \beta W(\eta, \theta) \geq u(\tau(\theta - \theta^*) \eta F(K)) + \beta U_0.$$

Proposition 2.7. *It is optimal to set the conditional base wage $\bar{w}(\eta) = \bar{w}$ and to set the termination wage $w_F(\eta) = \bar{w}$. The optimal choice of $\delta(\eta)$ is a simple cutoff rule where $\delta(\eta) = 1$ if $\eta \geq \bar{\eta}(U)$ and equal to 0 otherwise. The optimal cut-off level is decreasing in initial promised utility U .*

Proof: See the Appendix.

The result on the size of the insurance payment is an extension our prior results on compensations that compensation is constant at the base wage unless the no-perks constraint binds, and the base wage is independent of η . This constancy carries over to the case when the manager is being fired since compensation here is simply directed at his current flow utility. To understand this result, note once again that the marginal cost of a utility for a retained manager conditional on η is $\{1/u'(\bar{w}(\eta))\} \Pr\{w(\eta, \theta) = \bar{w}(\eta)|\eta\}$, while the benefit to the manager is $\Pr\{w(\eta, \theta) = \bar{w}(\eta)|\eta\}$. The marginal cost of utility for the manager when you fire him is $\{1/u'(w^F)\}$ while the conditional benefit is 1. Equating the cost-benefit ratios across these cases gives us simultaneously the constancy of the base wage at \bar{w} and the fact that $\bar{w} = w^F$.

The result that the cut-off is declining in U implies that the model exhibits a form of managerial entrenchment. Managers who have had better performance in the past will have higher continuation utilities, and these higher continuation utilities will make it more likely that the incumbent manager is retained in the future. Since managers who have been on the job longer will have a better chance of having had a high productivity shock $\theta > \bar{\theta}$ leading to a performance bonus and consequent increase in their future promised utility, managers

with greater tenure will on average be replaced less often than newer managers. The key aspect of the model that delivers our retention result is the fact that higher utility promises reduce the extent of agency frictions within the firm and thus make it cheaper to provide the incumbent manager with utility on the job than off it.

The predictions of our model are broadly consistent with the empirical findings in the literature. Our finding that low productivity shocks lead to managerial turnover is consistent with the empirical finding that executive turnover is correlated with poor performance as measured by either stock returns or accounting data, and that CEOs often receive large golden parachutes for leaving a firm in the wake of poor performance (see Kojima 1997 and Tirole 2006). Our findings are also consistent with the findings of Subramanian et al (2002) who find that CEOs with greater explicit incentives have less secure jobs, and those of Berger et al (1997) who find that leverage falls for a CEO with a long tenure, and weak stock and compensation incentive bonuses. Berger et al (1997) also find that the replacement of a long tenured CEO leads to an increase in leverage when the turnover appears "forced" (p.1436). As we already noted, our model implies that a CEO with longer tenures are more likely to have had past shocks which caused his no-perks constraint to bind, and hence have a high level of his base wage. Our comparative statics results imply (strictly speaking fixing K) that this high base wage will be associated with a lower level of monitoring and conditional on the level of monitoring, a decreased likelihood of his performance bonuses being triggered. Moreover, when the manager is replaced, the new manager will start at a lower utility promise and associated base wage, and hence the level of monitoring will be higher in this case (again, fixing K).

3 Interpreting the Optimal Contract

To interpret the other payments under this optimal contract in terms of debt and equity, we must ensure that payments to outside investors after the initial investment in the first sub-period are non-negative so that they do not violate the limited liability constraint imposed on investors in corporations. To do so, we assume that the outside investors invest not only the capital K , but also the noncontingent portion of the manager's pay \bar{w} in the first sub-period. We associate the payments v_0 or v_1 made by the manager in the second sub-period as the payments to debt holders. We associate the residual payments to outside investors as the payments to outside equity.

The payments made in the second sub-period are given by $v_1(\eta, \theta, \theta) = \theta\eta F(K)$ if $\theta \leq \theta^*(\eta)$ and $v_0(\eta, \theta) = \theta^*\eta F(K)$ if $\theta > \theta^*(\eta)$. We interpret $\theta^*\eta F(K)$ as the face value of the

project's debt. In the event that the realized value of the project exceeds the face value of the debt, the debt is paid. In the event that the realized value of the project is less than the face value of the debt, the project is bankrupt, monitored, and all remaining value is paid to the debt holders. If one assumes that the debt holders bear the cost of monitoring, the market value of a claim to the project's current debt payment is given by

$$D^A = \left[\left(\int_0^{\theta^*(\eta)} \theta p(\theta) d\theta + (1 - P(\theta^*(\eta))) \theta^*(\eta) \right) \eta - P(\theta^*) \gamma \right] \eta F(K(\eta)).$$

Note that under the assumption that the debt holders bear the cost of monitoring, the value of D can be negative since it is net of the cost of monitoring. Alternatively, one may assume that the outside investors jointly contribute resources $\gamma F(K)$ in addition to noncontingent payments K and \bar{w} in the first sub-period. Under this alternative assumption, the market value of a claim to the current debt payment is given by

$$D^B = \left[\left(\int_0^{\theta^*} \theta p(\theta) d\theta + (1 - P(\theta^*)) \theta^* \right) \eta + (1 - P(\theta^*)) \gamma \right] F(K),$$

which is always positive.¹⁰

The residual payout from the project is associated with the payments to the outside equity holders. In the event of bankruptcy ($\theta \leq \theta^*(\eta)$), the outside equity holders receive no payment. In the event that $\theta > \theta^*(\eta)$ the outside equity holders receive payment $(\theta - \theta^*) \eta F(K) - [w(\eta, \theta) - \bar{w}]$, which is the realized value of the project less the payment to the debt holders and the payments to the manager on the performance portion of his compensation. (Recall that the base portion of the manager's pay, \bar{w} , was set aside in advance). The value of a claim to the current payment is

$$E = \int_{\theta^*}^{\infty} [(\theta - \theta^*) \eta F(K) - w(\eta, \theta) + \bar{w}] p(\theta | \theta_{-1}) d\theta$$

There is an important issue that arises when one tries to determine the overall value of debt and equity claims on the firm. Note that the expected value of output less capital, compensation and bankruptcy costs,

$$\begin{aligned} & \int_{\theta} \{ \theta \eta F(K(\eta)) - w(\eta, \theta) \} p(\theta | \theta_{-1}) d\theta - P(\theta^* | \theta_{-1}) \gamma F(K(\eta)) - rK(\eta) \\ &= D^A + E - \bar{w} - rK(\eta), \end{aligned}$$

¹⁰This alternative assumption can also help rationalize commitment to deterministic monitoring since the proceeds from monitoring are nonnegative even if $\theta = 0$, and are positive for $\theta > 0$.

thus the value of debt and equity payments exceeds the value of net returns by the extent of the noncontingent claims $\bar{w} + rK(\eta)$. Even in a one period version of our model, this introduces an indeterminacy as to the initial value of these claims. If we assume that the initial owners of the firm assigned the responsibility for these noncontingent payments to the holders of equity claims, then the initial net value of equity is $E - \bar{w} - rK(\eta)$ and debt is D^A . Within a dynamic context, this would correspond to a situation in which debt holders had a long-term claim on payments D^A in every period, and the equity holders were received E less the payment of next period's noncontingent costs $\bar{w} + rK(\eta)$. In this case we would interpret debt as long-term bond with a coupon whose initial value was the present value of the stream of payments D^A . At the other extreme, assume that they assigned these costs to holders of the debt claim, in which case, the initial value of equity is E and debt is $D^A - \bar{w} - rK(\eta)$. Within a dynamic context, this would correspond to the case in which new one-period debt was issued in each period to cover the noncontingent costs, and the value of the long-term debt claim would be the present value of $D^A - \bar{w} - rK(\eta)$, while the value of equity would be the present value of the stream of payments E . Under different assumptions about the division of responsibility for ongoing investments in the firm, one obtains different implications for the debt-equity ratio of the firm. We conjecture that this issue will arise in any well-specified “trade-off” theory of optimal capital structure.

3.0.1 Capital Wedge

Condition (18) gives an analytic expression for the wedge between the internal and external rates of return on capital. To get a quantitative sense of the magnitude of this wedge, assume that we have log preferences, shocks are i.i.d., $\eta = 1$. In this case, the first-order condition with respect to capital becomes

$$\left\{ 1 - \gamma P(\theta^*) - \int_{\theta^*}^{\infty} \left[\frac{w(\theta) - \bar{w}}{F(K)} \right] p(\theta) d\theta \right\} F'(K) = r.$$

Bebchuk and Grinstein (2005) estimates the fraction of compensation paid by a large set of public firms to their top-five executives relative to net income at 8.1% over the 1999-2003 period. Over the same period, their average estimate of the share of equity-based compensation in total compensation at S&P 500 firms is 65%. This implies a wedge of roughly 5.3% from the compensation factor alone.

3.0.2 Risk Hedges and Public Signals

Financial hedges are contracts that firms enter into in order to insure themselves against certain (typically) exogenous events. Why do we see firms using financial hedges? The standard Modigliani-Miller logic would suggest that they have no role to play. In the literature, it has been argued that these financial hedges can be used to avoid risks which can lead to bankruptcy (Smith and Stoltz 1985) or to reduce the risk associated with stochastic cash-flows when external funds are more costly than internal funds (Froot, Scharfstein and Stein 1993).¹¹ Our models suggests a very different motivation. While our optimal contracting problem is sufficiently general to allow the firm to hedge risks, our results indicate that there is an efficient contract without such hedges. The result that financial hedges do not add value in our basic model emerges because the debt and equity contracts have been optimally chosen to offset the enforcement and incentive problems the outside investors fare with respect to the manager. We see, in particular, that in the second sub-period, the outside investors want to extract from the firm as large a payment as is possible given the choice of monitoring. Additional funds paid into the firm at this point would only exacerbate the agency friction as modelled by the no-perks constraint.

However, if we alter our model by assuming that η is observed at the beginning of the second subperiod, it now becomes an informative public signal of the firm's second subperiod. This change will lead to the capital choice being independent of η and the first-order condition with respect to K now including the integral not only over θ but also over η . However, it would still be possible to condition the monitoring decision on η , and the first-order condition for monitoring would be unchanged; modulo the replace of $K(\eta)$ with K . In this case, the optimal contract would require that the monitoring threshold still depend upon the realized η .

One way to implement the efficient level of monitoring would be with state-contingency as to the face value of the debt. However, one can also implement this state-contingency with non-contingent debt and financial hedges. To see how this is done, take $\theta^*(\eta)$ as the optimal monitoring threshold, and take θ^D as our noncontingent debt payment that is required to avoid monitoring, where $\theta^D : \theta^*(\eta^l) < \theta^D < \theta^*(\eta^h)$. Then we need to have a security with payoff $(\theta^*(\eta) - \theta^D) \eta F(K)$. This implies that the financial hedge will take the form of insurance against η realizations. Thus, the hedge is simultaneously smoothing the net income of the firm and reducing the sensitivity of managerial compensation to the unobserved component of output shocks, i.e. θ .

¹¹Acharya and Bisin (2005) have recently argued the hedges can be use to reduce the incentive of risk averse managers to skew investment choices towards projects with aggregate risk that they can more readily offset in their private portfolios than idiosyncratic risk.

In this theory, the purpose of the hedge is not to remove or reduce the risk of bankruptcy with simple debt contracts, but rather to fine tune it to allow the monitoring associated with bankruptcy to be undertaken in the optimal state-contingent fashion. One advantage of this approach may be that rather than having the firm market a unique type of state-contingent debt security; it instead markets a standard debt security and, assuming it's available, acquires a set of positions in a standard financial hedging contract.

3.0.3 Salvage Option

Now consider the interpretation of monitoring in our dynamic model. In interpreting our efficient contract as a theory of capital structure, we associate monitoring with bankruptcy. Monitoring in our model occurs whenever the current gross output of the firm fall below a threshold $\theta^*(\eta)\eta F(K(\eta))$ determined by the optimal contract. In the event that $\theta \leq \theta^*(\eta)$, monitoring occurs, but the firm still has a value to the outside investors as an ongoing concern (denoted by the continuation value $V(W(\theta))$). In the event that this continuation value exceeds the face value of the debt, then the equity holders emerge from this episode of bankruptcy with shares that still have positive value. In this sense, monitoring in the dynamic model does not necessarily correspond to the liquidation of the firm. Of course, the same is true of bankruptcy in the data.

Alternatively, we could have assumed that monitoring destroyed the firm, and that it had a salvage value S . In this case, the continuation payoff to the investors would become

$$I(\theta > \theta^*(\eta))\frac{1}{R}V(W(\eta, \theta), \eta, \theta) + I(\theta \leq \theta^*(\eta))S,$$

where I denotes an indicator function. Assuming that the continuation value of the agent was simply U_0 when the firm ceased to exist, the continuation payoff to the manager would be given by

$$I(\theta > \theta^*(\eta))W(\eta, \theta) + I(\theta \leq \theta^*(\eta))U_0,$$

and the no-perks constraint would continue essentially unchanged. To the extent that continuation values exceed the salvage cost, this would make monitoring more costly, but would leave the essential characterization of the optimal contract unchanged, except that it would introduce the issue of compensation in the case of termination.

3.0.4 Performance Bonus

The exact features of the performance bonus schedule predicted by our model depends on several factors. First, if $\beta = 0$, then $\tilde{w}(\theta) = \bar{w} + \tau(\theta - \bar{\theta})F(K)$, and the performance bonus

is linear in the output of the firm net of the face value of its debt. However, in general the current performance bonus increases by less than one-to-one with τ times net output both because the bonus is smoothed over time, with a bonus today being associated with a higher base wage tomorrow, and because this smoothing reduces the present value of the total payment to the manager needed to offset what he can grab.

Second, in our model, the performance component of the manager's is triggered by the value of the output of the firm relative to the face value of its maturing debt, $\tau(\theta - \theta^*(\eta))\eta F(K(\eta))$. This implication is driven by the exact form of our agency friction. To see this, consider a variant of our model in which firm output had two components: current cash flow $\theta\eta f(K)$ and undepreciated capital $(1 - \delta)K$. Assume that the manager is able to spend up to fraction τ of undisbursed cash flow on perquisites, but that he cannot divert undepreciated capital for his own use. In this variant of the model, the constraint on payments to the manager would be modified to be

$$u\left(x(\eta, \hat{\theta}, \theta)\right) + \beta Z(\eta, \hat{\theta}, \theta) \geq u(\tau(\theta\eta f(K) - v_0(\eta, \hat{\theta}))) + \beta U_0$$

if monitoring did not occur, and

$$u\left(x(\eta, \hat{\theta}, \theta)\right) + \beta Z(\eta, \hat{\theta}, \theta) \geq u(\tau(\theta\eta f(K) - v_1(\eta, \hat{\theta}, \theta))) + \beta U_0$$

otherwise. The limited liability constraint would be modified to read

$$v_0(\eta, \hat{\theta}) \leq \hat{\theta}\eta f(K) + (1 - \delta)K, \quad v_1(\eta, \hat{\theta}, \theta) \leq \theta\eta f(K) + (1 - \delta)K, \quad \text{and } x(\eta, \hat{\theta}, \theta) \geq 0.$$

It is straightforward to show that the optimal contract in this variant of the model would break down into four payments as before, except in this case, the performance pay to the manager would be based on cash flow $\theta\eta f(K)$ and not on the value of the firm (which includes the value of undepreciated capital). It is also straightforward in this variant of the model to interpret the payments v backed by undepreciated capital $(1 - \delta)K$ as payments to collateralized debt.

Finally, it is also worth noting that the shape of the response of the performance bonus is sensitive to our assumption that the manager can steal a constant fraction τ of residual output. Nothing in our qualitative results would be changed if we assumed that the amount he can steal was an increasing function of residual output. Of course the shape of this function and whether it was concave or convex would have an important impact on the response of performance bonus to increases in θ .

4 Concluding Comments

This paper presents a model of capital structure and executive compensation based upon two frictions internal to the firm: an information friction and an agency friction. These frictions are potentially binding for the duration of an information cycle. The frictions motivate the division of firm's payout into debt and equity payments, and the division of compensation into base pay, a performance bonus, and a golden parachute style severance package for managers who are terminated because of insufficient productivity. We show how to collapse these two frictions into a single no-perks constraint. We found that the extent to which the no-perks constraint binds determines the extent of monitoring and hence the capital structure of the firm. It also determined the wedge between the internal and external return to capital and the share of executive compensation coming from performance pay.

In our model, limitations on ex post punishments along with the competing desires to smooth compensation to the manager in order to reduce the cost of his consumption, and to backload his compensation in order to reduce the extent to which the no-perks constraint binds generates a very stark connection between current compensation and the next period's base wage. This connection meant that shocks that lead to the no-perks constraint binding today, and hence the payment of a performance bonus today, also lead to an increase in the future base wage. This increase in the future base wage causes a reduction in the future extent to which the no-perks constraint binds, which in turn impacts on the future capital structure of the firm and the future compensation scheme of the manager. We also show how the growth prospects of the firm, or the persist effects of shocks, impact on the current financial structure and compensation scheme through the continuation payoff of the manager. Factors that increase his continuation payoff lead to a reduction in the extent to which our no-perks constraint binds today and hence a reduction in monitoring and the extent to which the manager's compensation came in the form of a performance bonus.

Many of the model's predictions are consistent with the empirical literature on executive compensation and capital structure. Just as in our model, executive compensation is downwardly rigid, luck is rewarded, and relative performance is not a factor. The model's predictions that poor performance leads to managerial turnover, that managers whose compensation is more heavily weighted towards performance bonuses have less secure jobs and their firms leverage ratios are higher are also consistent with the data. The model's theory of capital structure predicts that size will be correlated with leverage, and that the market-to-book ratio of asset values will be negatively with leverage, just as it is in the data.

Several surprising findings came out of our analysis. Bankruptcy emerges as means of achieving optimal monitoring, not because of solvency. Managerial entrenchment can

be efficient since it is cheaper to compensate managers within the firm than via a golden parachute. Hedging turns out to achieve efficient trade-off between bankruptcy risk and agency risk with nonstate-contingent debt, rather than as a way to reduce bankruptcy risk.

References

- [1] Aghion, P. and P. Bolton, 1992. An Incomplete Contracts Approach to Financial Contracting, *Review of Economic Studies*, 59(3): 473-94.
- [2] Albuquerque, R. and Hopenhayn, H. "Optimal Lending Contracts and Firm Dynamics", with Hugo Hopenhayn, September 2002. Forthcoming *Review of Economic Studies*.
- [3] Alvarez, F. and U. Jermann, 2000. Efficiency, Equilibrium, and Asset Pricing with Risk of Default, *Econometrica*, 68(4): 775-97.
- [4] Aiyagari, S. Rao, and Fernando Alvarez. "Stationary Efficient Distributions With Private Information and Monitoring: A Tale of Kings and Slaves." Manuscript, Federal Reserve Bank of Minneapolis, 1995.
- [5] Atkeson, A., "International Lending with Moral Hazard and Risk of Repudiation," *Econometrica*, 79:14-31, 1991.
- [6] Atkeson, A. and H. Cole. "A Dynamic Theory of Optimal Capital Structure and Executive Compensation," National Bureau of Economic Research Working Paper #11083.
- [7] Acharya, V. and A. Bisin 2005. Managerial Hedging, Equity Ownership, and Firm Value, N.Y.U. research memo.
- [8] Berger, P. E. Ofek, D. Yermack, 1997, Managerial Entrenchment and Capital Structure Decisions. *Journal of Finance* 52(4) 1411-1438.
- [9] Bernanke, B. and M. Gertler 1989. Agency Costs, Net Worth, and Business Fluctuations, *American Economic Review*, 79(1), 14-31.
- [10] Carlstrom C. and T. Furest 1997. Agency Costs, Net Worth, and Business Fluctuations: A Computable General Equilibrium Model, *American Economic Review*, vol 87(5), 893-910.
- [11] Clementi, G.L. and Hopenhayn, H., "A Theory of Financing Constraints and Firm Dynamics," RCER working paper #492, June 2002.

- [12] Cooley, T., R. Marimon and V. Quadrini, “Aggregate Consequences of Limited Contract Enforceability,” *Journal of Political Economy*, 112(4): 817-847.
- [13] DeMarzo, P. and Fishman, M., “Optimal Long-Term Financial Contracting with Privately Observed Cash Flows,” Graduate School of Business, Stanford University, 2003.
- [14] Dewatripont, M. and J. Tirole. 1994. A Theory of Debt and Equity: Diversity of Securities and Manager-Shareholder Congruence, *Quarterly Journal of Economics*, 109(4): 1027-54.
- [15] Froot, K. , D. Scharfstein, and J. Stein. 1993. Risk Management: Coordinating Corporate Investment and Financing Policies. *Journal of Finance*, vol. 48 (5), 1629-58.
- [16] Green, E. “Lending and the Smoothing of Uninsurable Income.” In Edward C. Prescott and Neil Wallace, eds., *Contractual Arrangements for Intertemporal Trade*, pp. 3-25. Minneapolis: University of Minnesota Press, 1987.
- [17] Hart, O. and J. Moore. 1995. Debt and Seniority: An Analysis of the Role of Hard Claims in Constraining Management, *American Economic Review*, 85(3): 567-85.
- [18] Hart, O. and J. Moore, 1998. Default and Renegotiation: A Dynamic Model of Debt, *Quarterly Journal of Economics*, 113(1), 1-41.
- [19] Jensen, M. 1986. Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers, *American Economic Review*, 76(2): 323-29.
- [20] Kehoe, T. and D. Levine, 1993. Debt-Constrained Asset Markets, *Review of Economic Studies*, 60(4): 865-88.
- [21] Kojima, K. 1997. Corporate Governance: An International Comparison, Hajime Printing.
- [22] Myers, S.C., 1984. The Capital Structure Puzzle, *Journal of Finance* 39(July): 572-592
- [23] Modigliani, F. and M. H. Miller, 1958. The Cost of Capital, Corporate Finance and the Theory of Investment, *American Economic Review*, 48(June): 261-297
- [24] Quadrini, V., “Investment and Default in Optimal Financial Contracts with Repeated Moral Hazard,” Stern School of Business, New York University, 2000.
- [25] Rajan, R. and L. Zingales, 1995. What Do We Know about Capital Structure? Some Evidence from International Data, *Journal of Finance*, 50(5), 1421-1460.

- [26] Smith, C. and R. Stoltz 1985. The Determinants of Firm's Hedging Policies, *Journal of Financial and Quantitative Analysis* 20, 391-405.
- [27] Subramanian, N., A. Chakraborty, and S. Sheikh, 2002. Performance Incentives, Performance Pressure and Executive Turnover, research mimeo, Brandis University.
- [28] Tirole, J. *The Theory of Corporate Finance*, Princeton University Press, 2006.
- [29] Townsend, R. 1979. Optimal Contracts and Competitive Markets with Costly State Verification, *Journal of Economic Theory*, 21(2): 265-93.
- [30] Wang, C. , 2004. Dynamic Costly State Verification, research memo.
- [31] Welch, I. 2004. Capital Structure and Stock Returns, *Journal of Political Economy*, 112(1), 106-131.
- [32] Zweibel, J. 1996. Dynamic Capital Structure under Managerial Entrenchment, *American Economic Review*, 86 (5), 1197-1215.

5 Appendix

Proof of Proposition 2.5 Fixing $K(\eta)$, we can show that $d\theta^*(\eta)/d\bar{w} < 0$ by the following argument: (22) is the f.o.c. w.r.t. $\theta^*(\eta)$. If we are an interior maximum, then it must be the case that the derivative of l.h.s. minus the r.h.s. with respect to $\theta^*(\eta)$ is negative. The derivative of (22) w.r.t. \bar{w} is negative. To see this note that for θ such that $\max[\bar{w}, \tilde{w}(\eta, \theta)] = \bar{w}$, the value inside the integral is zero. Then note that for those θ such $\max[\bar{w}, \tilde{w}(\eta, \theta)] > \bar{w}$, $d \max[\bar{w}, \tilde{w}(\eta, \theta)] / d\bar{w} = 0$. Hence, because the derivative of $-1/u'(\bar{w})$ w.r.t. \bar{w} is negative and the result follows.

We can show that $d\theta^*(\eta)/d\gamma < 0$ by simply noting that the derivative of the first-order condition for $\theta^*(\eta)$ w.r.t. γ is negative. Since, as we have already noted, the second derivative of the Lagrangean with respect to $\theta^*(\eta)$ is negative, the results follows.

Fixing $\theta^*(\eta)$, we can show that $dK(\eta)/d\bar{w} > 0$ by the following argument: (23) is the f.o.c. w.r.t. $K(\eta)$, and hence the second derivative of the l.h.s. is negative at an interior optimum. The same argument as before implies that the derivative of r.h.s. w.r.t. \bar{w} is positive here because of the negative sign in front of the integral. *Q.E.D.*

Discussion of Proposition: To understand why we cannot get an overall result, assume that preferences are CRRA, and note first that if $\beta = 0$, then

$$\tilde{w}(\eta, \theta) = \tau(\theta - \theta^*(\eta))\eta F(K(\eta)),$$

and,

$$\begin{aligned} & \frac{u' [\tau(\theta - \theta^*(\eta))\eta F(K(\eta))]}{u' (w(\eta, \theta))} - \frac{u' [\tau(\theta - \theta^*)\eta F(K(\eta))]}{u'(\bar{w})} \\ = & 1 - \frac{u' [\tau(\theta - \theta^*)\eta F(K(\eta))]}{u'(\bar{w})}, \end{aligned}$$

in which case it's straightforward to show that $d\theta^*(\eta)/d(F(K)) > 0$. As β goes to 1, then $\tilde{w}(\eta, \theta)$ doesn't respond to the increase in $F(K(\eta))$, which implies that

$$\begin{aligned} & \frac{u' [\tau(\theta - \theta^*(\eta))\eta F(K(\eta))]}{u' (w(\eta, \theta))} - \frac{u' [\tau(\theta - \theta^*)\eta F(K(\eta))]}{u'(\bar{w})} \\ \simeq & \left(\frac{u' [\tau(\theta - \theta^*(\eta))\eta]}{u' (w(\eta, \theta))} - \frac{u' [\tau(\theta - \theta^*)\eta]}{u'(\bar{w})} \right) u'(F(K(\eta))), \end{aligned}$$

and hence it will follow that $d\theta^*(\eta)/d(F(K)) < 0$. When $d\theta^*/dK > 0$ an increase in \bar{w} is having two effects: (i) a direct effect which tends to lower θ^* , and (ii) an indirect effect through a potential increase in K coming from the increase in \bar{w} , which tends to raise θ^* . These offsetting effects also make it difficult to derive general results also with respect to how θ^* and K vary with η .

Proof of Proposition 2.7 The fact that $\bar{w}(\eta) = \bar{w}$ follows trivially from the same argument as in proposition 2.2. Given this, it follows that $w_F = \bar{w}$.

Assume that $\bar{\eta}$ was such that

$$\begin{aligned} & V^R(U^R(\bar{\eta}), \bar{\eta}) + \omega U^R(\bar{\eta}) \\ = & V^R(U_0, 1) - w^F(\bar{\eta}) + \omega [u(w^F(\bar{\eta})) + \beta U_0], \end{aligned}$$

or another words the principal was just indifferent between retaining and firing the manager. Then, consider differentiating both sides w.r.t. η , given the optimum choices $U^R(\eta)$ and $w^F(\eta)$. Note the derivative of the r.h.s. is 0 since w^F is independent of η , while the derivative of the l.h.s. is

$$-\omega U^{R'}(\bar{\eta}) + D_2 V^R(U^R(\bar{\eta}), \bar{\eta}) + \omega U^{R'}(\bar{\eta}) = D_2 V^R(U^R(\bar{\eta}), \bar{\eta}).$$

To see that $D_2 V^R(U^R(\eta), \eta) > 0$, note that the contract could always offset the impact of η on $\eta F(K)$ by lowering K . This reduces the cost of capital but otherwise leaves the second state problem unchanged. Hence the result follows. *Q.E.D.*

Figure 1: No Public Shocks

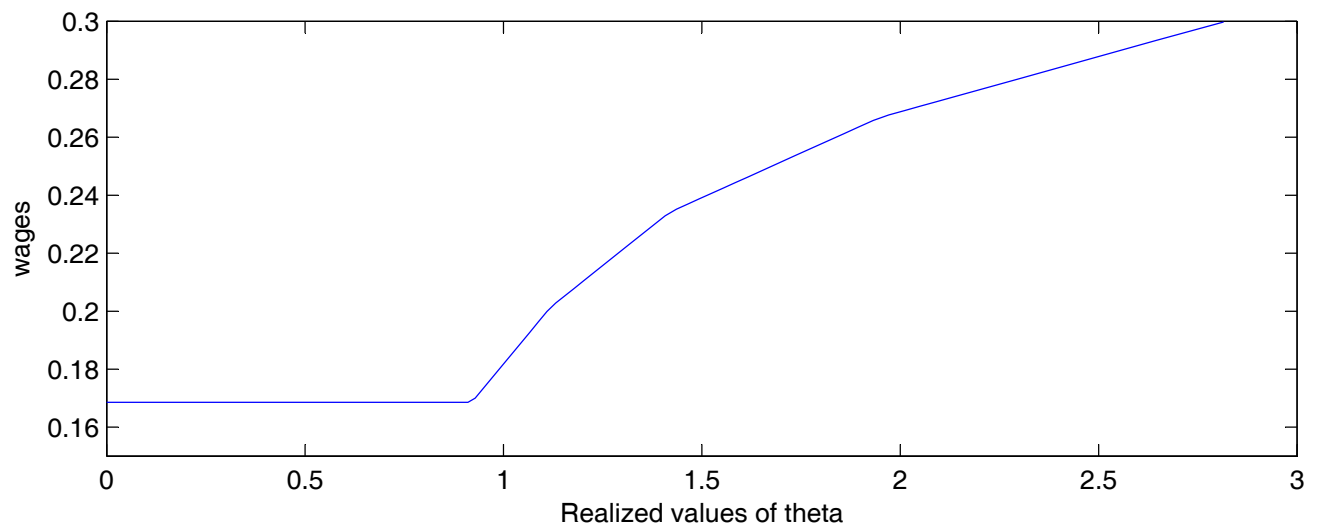
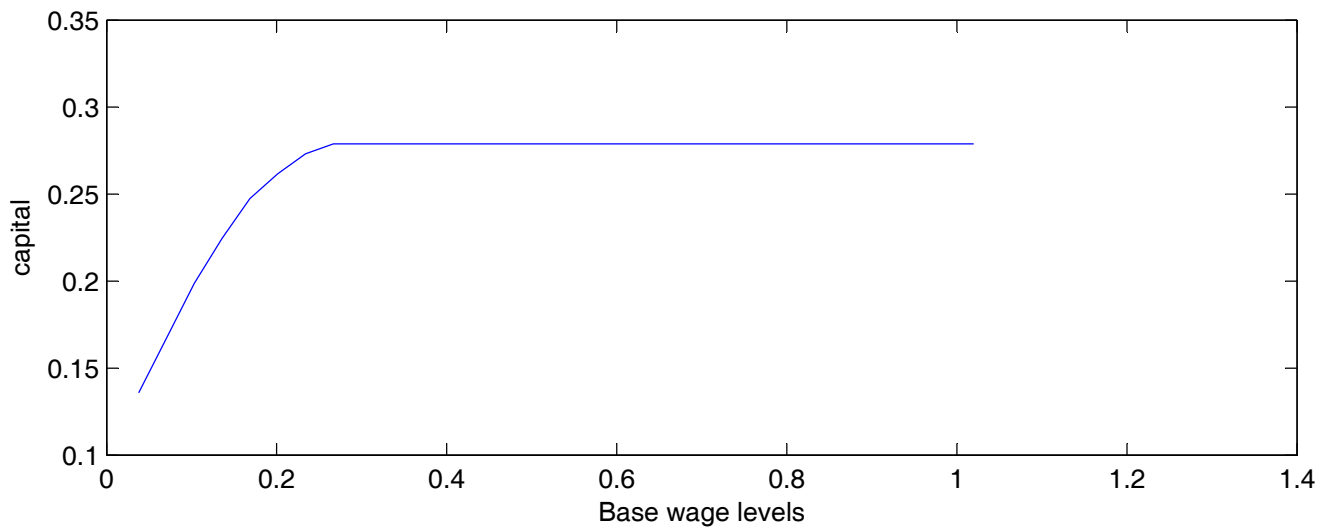
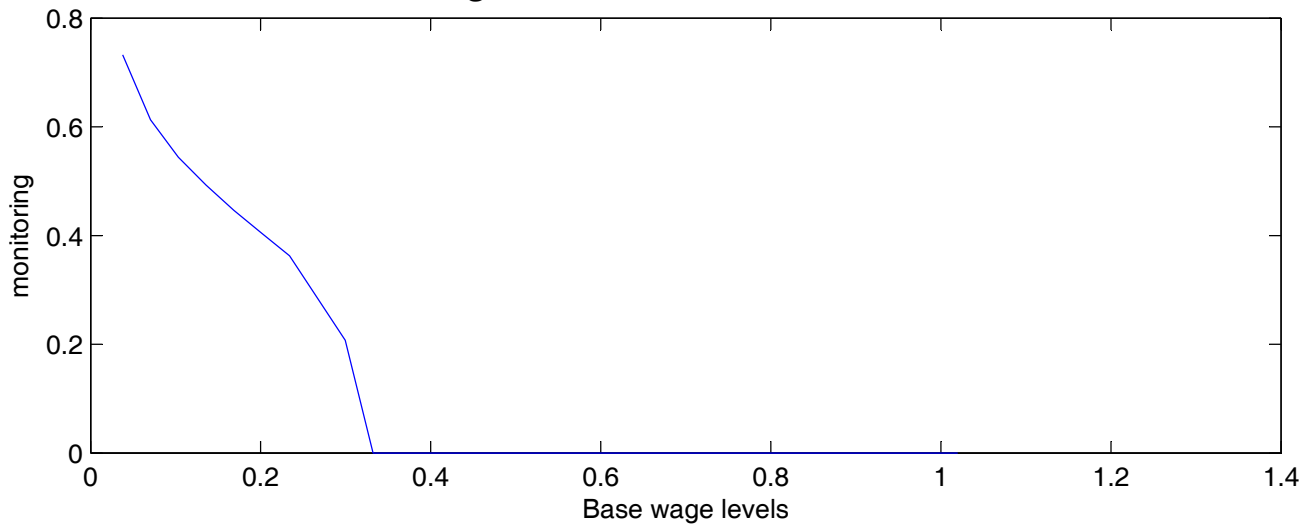


Figure 2: Growth vs. No Growth

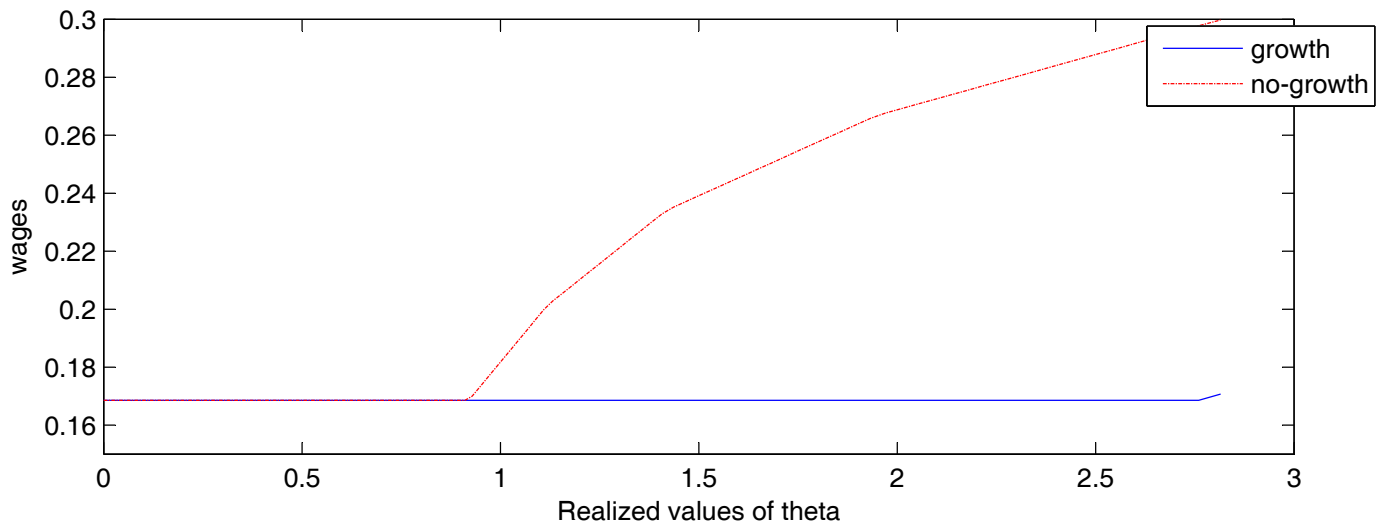
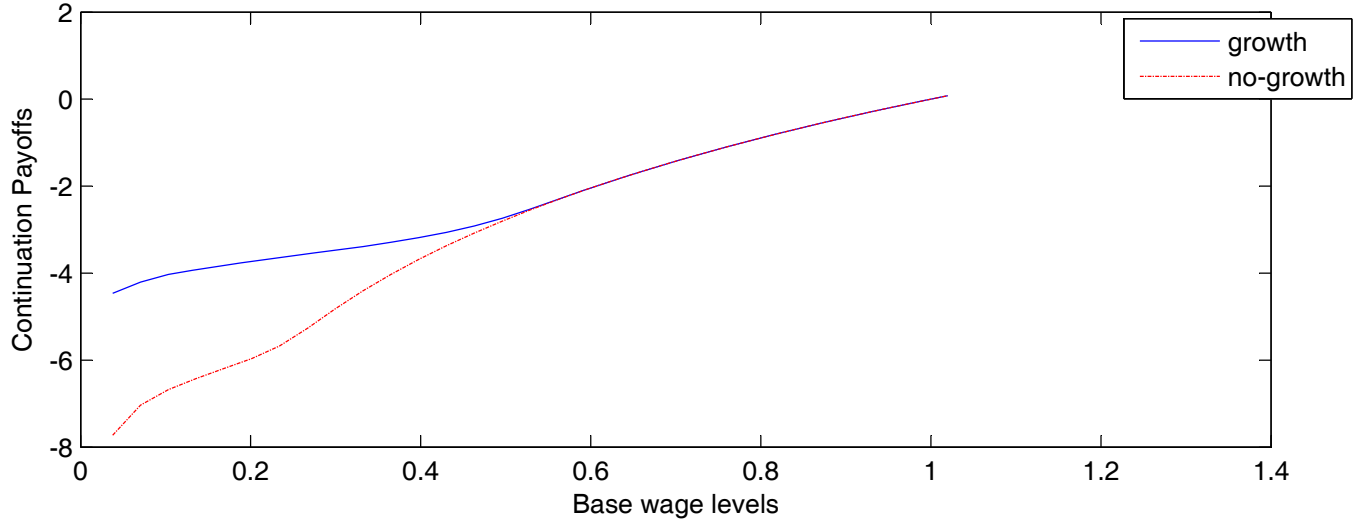
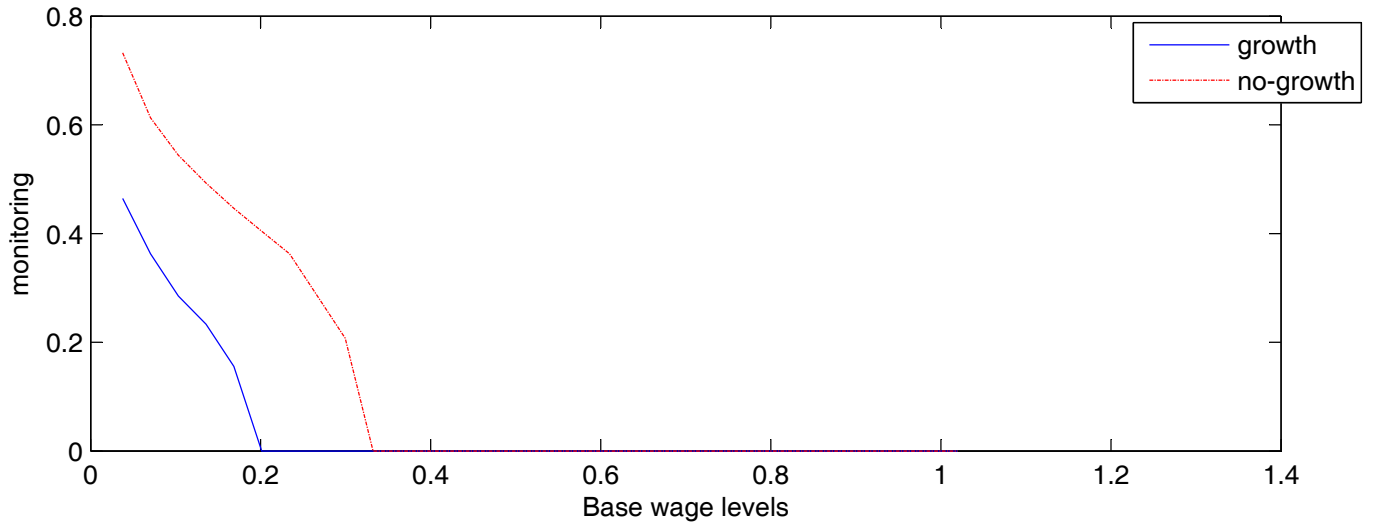


Figure 3: IID Case

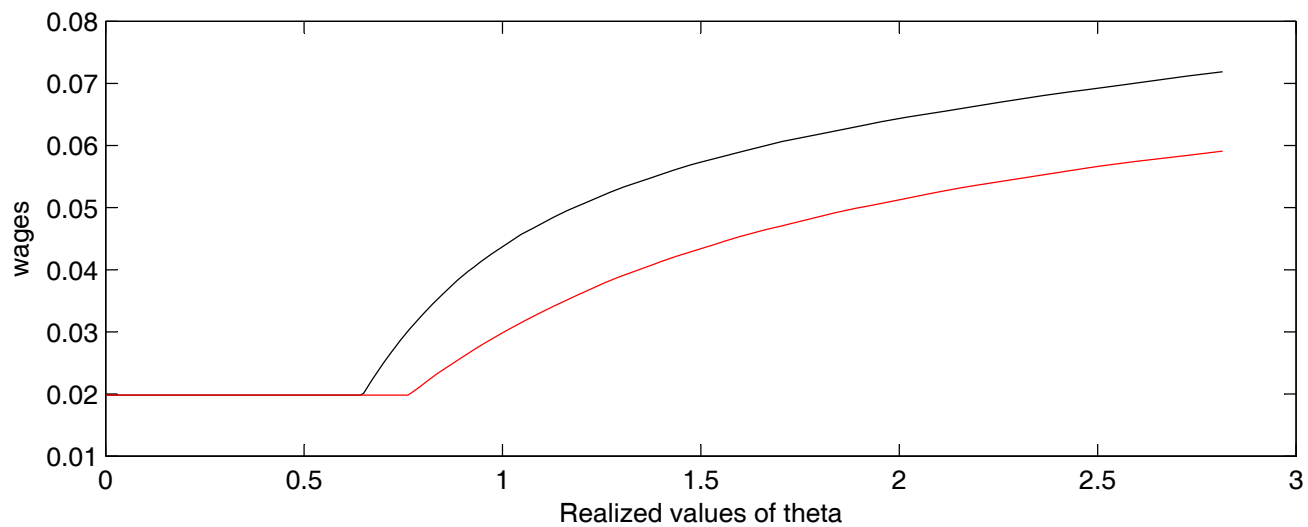
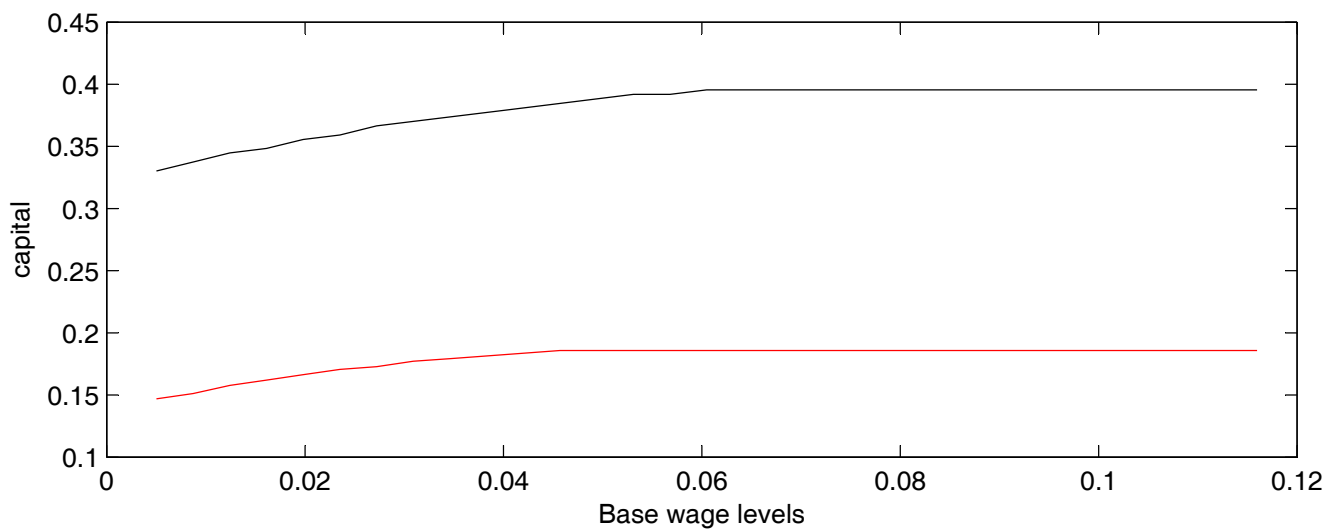
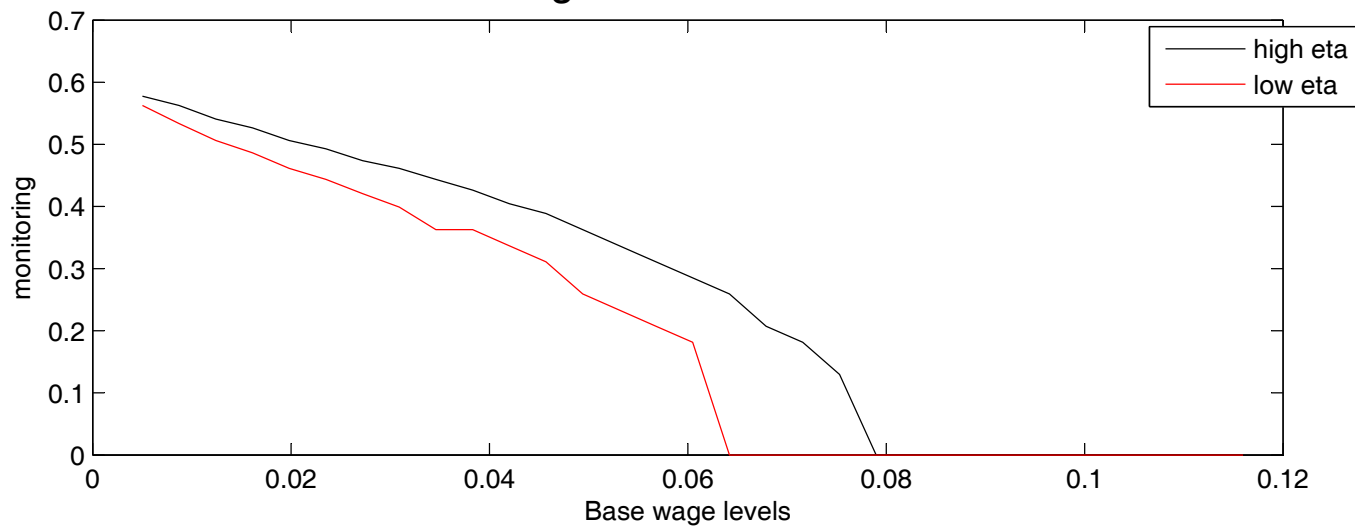


Figure 4: Comparing the IID and Persistent Cases

